NATIONAL RESEARCH UNIVERSITY

HIGHER SCHOOL OF ECONOMICS

International College of Economics and Finance

Grigorii Kuzmin

gikuzmin@edu.hse.ru

Evaluating the Adjusted PIN Model (Duarte and Young, 2007) Efficiency of Measuring the Probability of Informed Trading in Cryptocurrency Markets

Msc Coursework

38.04.01 ECONOMICS

Master's Programme

"HSE master's degree programme in financial economics"

Scientific advisor

PhD

Alexei Boulatov

Moscow 2022

Introduction

Evaluation of informed trading has been one of the most topical research directions in financial microstructure for the last decades. One of the most popular frameworks is PIN model, introduced by Easley et al (1996), which uses the Glosten Milgrom (1985) framework to model the arrival of private signals to the market and the subsequent effect on trading rates of different types of traders. However, this model suffered from several inefficiencies: floating point exception, corrected by Lin and Ke (2011), the maximum likelihood maximization issues, related to finding local but not the global maximum, solved by Yan and Zhang (2012). However, apart from these computational problems, this model failed to explain the empirically observed positive correlation between Buy and Sell trades as it mathematically does not allow for such relation sign. Thus, Duarte et al (2007) introduced the Adjusted PIN, which is the extension of the traditional model with additional parameters, allowing for simultaneous Buy and Sell operations event of informed traders.

One the main drawbacks of traditional PIN model, mentioned above, was the so-called **Floating Point Error (FPE)** problem which means the computer receives a value exceeding its computational capacity. In other words, modern computational software is unable to estimate $exp(x)$ where $x > 710$. This problem was solved by Lin and Ke (2011) who transformed PIN Likelihood function in such a way that guarantees its values will not exceed this boundary during maximization process.

Another weakness was related to the maximization process as very often software algorithms provide boundary solutions or stop at local maximum, but not the global one. Yan and Zhang (2012) derived an initial parameters algorithm which creates 125 initial starting points for optimization problem, increasing the probability of find the global maximum and helping to avoid **false boundary solutions**.

Adjusted PIN by Duarte et al (2007) is a nested version of PIN, so it inherits both **FPE** and **boundary values** problems. However, no approaches to solve this issues havr been introduced officially yet. Thus, the main aim of this paper is, using the ideas of the two aforementioned techniques, derive new supporting algorithms for Adjusted PIN model to solve these two computational inefficiencies.

The paper is organized in the following fashion: in *Section 1* we provide a brief description of the Adjusted PIN model and theoretical review of other models implied, in *Section 2* we provide a new version of the Likelihood function, and derive a new version of Initial Parameters algorithm, while in *Section 3* data generation process is described and, finally, in *Section 4* we prove the efficiency of algorithms proposed, based on the simulated data.

1 Literature Review [1]

Adjusted PIN

This model, introduced by Duarte and Young (2007) is an extension of PIN (EHO) model. As in the original model, there are two types of traders (insiders and noise), and with probability α there is a private information event, which can be either positive or negative with underlying probabilities δ and $(1-\delta)$, respectively. However, we allow informed traders to be heterogeneous and perform Buy/Sell trades at different rates: μ_b and μ_s. Moreover, we introduce an event of symmetric order flow, leading to additional Sell (Δ_s) and Buy (Δ_b) orders at the same time. This modification enables the model to match the empirically observed positive correlation between Buy and Sell orders, which the traditional PIN model fails to do.

[1] Partly based on the bachelor thesis paper G.Kuzmin (2021)

Figure 1: Adjusted PIN model Prob. tree

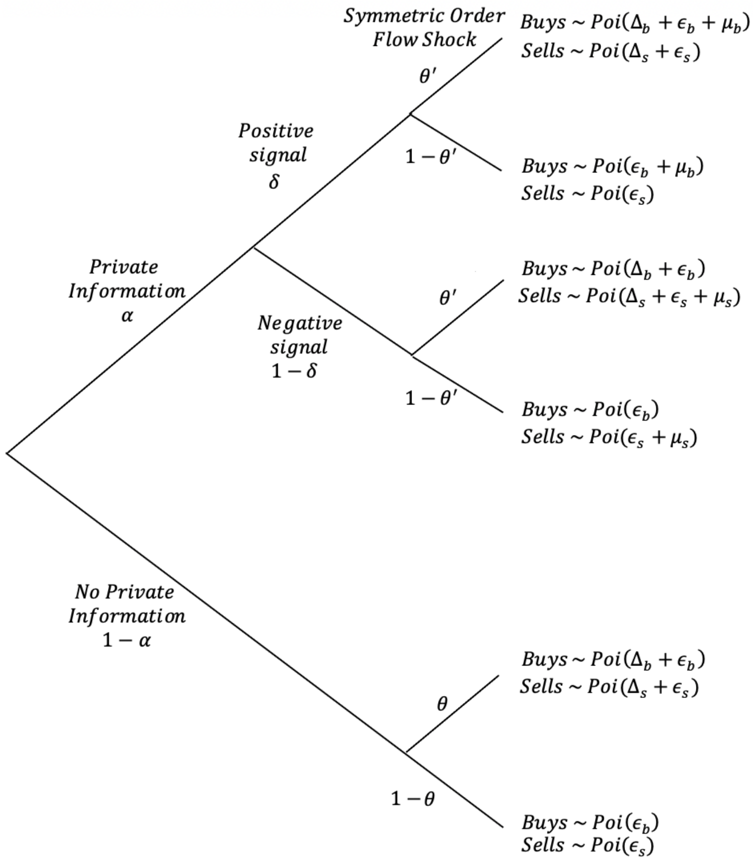


Figure 1: **Adjusted trading process tree.** This diagram represents the trading mechanics, where α, δ, θ, θ', μ_b, μ_s, ϵ_b, ϵ_s, Δ_b and Δ_s stay for probabilities of private information event, of positive signal and of symmetric order flow in case of private event and its absence, rate of informed buy and sell operations, rates of noisy buy and sell operations and symmetric buy and sell rates, respectively.

Using the Poisson distribution assumption, we arrive at the following Likelihood function:

$$
\begin{aligned}
L(\Theta|B_i,S_i) = ln \Bigg[& (1-\alpha)(1-\theta)exp(-\epsilon_b-\epsilon_s)\frac{\epsilon_b^{B_i}\epsilon_s^{S_i}}{B_i!S_i!}+ \\
& + (1-\alpha)\theta exp(-\epsilon_b-\epsilon_s-\Delta_b-\Delta_s)\frac{(\epsilon_b+\Delta_b)^{B_i}(\epsilon_s+\Delta_s)^{S_i}}{B_i!S_i!}+ \\
& + \alpha(1-\theta')(1-\delta)exp(-\epsilon_b-\mu_s-\epsilon_s)\frac{\epsilon_b^{B_i}(\mu_s+\epsilon_s)^{S_i}}{B_i!S_i!}+ \\
& + \alpha\theta'(1-\delta)exp(-\epsilon_b-\epsilon_s-\mu_s-\Delta_b-\Delta_s)\frac{(\epsilon_b+\Delta_b)^{B_i}(\mu_s+\epsilon_s+\Delta_s)^{S_i}}{B_i!S_i!}+ \\
& + \alpha(1-\theta')\delta exp(-\mu_b-\epsilon_b-\epsilon_s)\frac{(\mu_b+\epsilon_b)^{B_i}\epsilon_s^{S_i}}{B_i!S_i!}+ \\
& +\alpha\theta'\delta exp(-\mu_b-\epsilon_b-\epsilon_s-\Delta_b-\Delta_s)\frac{(\mu_b+\epsilon_b+\Delta_b)^{B_i}(\epsilon_s+\Delta_s)^{S_i}}{B_i!S_i!}\Bigg]
\end{aligned}
$$

where $\Theta = (\alpha,\delta,\theta,\theta^{'},\mu_b,\mu_s,\epsilon_b,\epsilon_s,\Delta_b,\Delta_s)$ are no news, good news, probability of symmetric buy and sell trades, given there is NO private signal, probability of symmetric buy and sell trades, given there is private signal, insider's buy and sell trading rates, noise traders' buy and sell trading rates, additional buy and sell trading rates in case of symmetric trading event, respectively, while B and S are total Buy and Sell operations per day.

In order to tackle factorials, which cannot be computed for large numbers, Duarte et al (2007) modifies the Likelihood, using $e^{-\epsilon_b}\frac{\epsilon_b^B}{B!} \approx e^{-\epsilon_b+B\cdot ln(\epsilon_b)-\sum_{i=1}^{B} ln(i)}$

Thus, we introduce Duarte Likelihood:

$$
\begin{aligned}
L(\Theta|B,S) =& (1-\alpha)(1-\theta)e^{-\epsilon_b}\frac{\epsilon_b^B}{B!}e^{-\epsilon_b}\frac{\epsilon_s^S}{S!} \\
&+(1-\alpha)\theta e^{-(\epsilon_b+\Delta_b)}\frac{(\epsilon_b+\Delta_b)^B}{B!}e^{-(\epsilon_s+\Delta_s)}\frac{(\epsilon_s+\Delta_s)}{S!} \\
&+\alpha(1-\theta')(1-\delta)e^{-\epsilon_b}\frac{\epsilon_b^B}{B!}e^{-(\epsilon_s+\Delta_s)}\frac{(\epsilon_s+\Delta_s)^S}{S!} \\
&+\alpha\theta'(1-\delta)e^{-(\epsilon_b+\Delta_b)}\frac{(\epsilon_b+\Delta_b)^B}{B!}e^{-(\mu_s+\epsilon_s+\Delta_s)}\frac{(\mu_s+\epsilon_s+\Delta_s)^S}{S!} \\
&+\alpha(1-\theta')\delta e^{-(\mu_b+\epsilon_b)}\frac{(\mu_b+\epsilon_b)^B}{B!}e^{-\epsilon_s}\frac{\epsilon_s^S}{S!} \\
&+\alpha\theta'\delta e^{-(\mu_b+\epsilon_b+\Delta_b)}\frac{(\mu_b+\epsilon_b+\Delta_b)^B}{B!}e^{-(\epsilon_s+\Delta_s)}\frac{(\epsilon_s+\Delta_s)^S}{S!}
\end{aligned}
$$

Using independence of information signals for each particular day, we can reformulate the maximization problem for t periods as:

$$V = \prod L(\Theta|B,S) = \sum logL(\Theta|B,S)$$

Formula for Adjusted PIN is the ratio of expected insider trading order flow to total order flow (nested PIN formula):

$$Adj\ PIN = \frac{\alpha \times (\delta \times \mu_s + (1-\delta) \times \mu_b)}{\alpha \times ((1-\delta) \times \mu_b + \delta \times \mu_s) + (\Delta_b + \Delta_s) \times (\alpha \times \theta' + (1-\alpha) \times \theta) + \epsilon_s + \epsilon_b}$$

Lin and Ke (2011) factorization

To solve **FPE** problem in PIN Likelihood function, Lin and Ke (2011) inroduced one od the most efficient and widely used factorization techniques, producing the following modified PIN Likelihood:

$$
\begin{aligned}
L(\Theta|B_t,S_t) = & \ln\left[\alpha\delta exp(e_{1i}-e_{maxi}) + \alpha(1-\delta)exp(e_{2i}-e_{maxi}) + (1-\alpha)exp(e_{3i}-e_{maxi})\right]+ \\
& + B_t\ln(\epsilon_b+\mu) + S_t\ln(\epsilon_s+\mu) - (\epsilon_b+\epsilon_s) + e_{maxi} - \ln(S_t!B_t!) \quad (1)
\end{aligned}
$$

where $e_{1i} = -\mu - B_t \ln(1 + \frac{\mu}{\epsilon_b})$, $e_{2i} = -\mu - S_t \ln(1 + \frac{\mu}{\epsilon_s})$, $e_{3i} = -B_t \ln(1 + \frac{\mu}{\epsilon_b}) - S_t \ln(1 + \frac{\mu}{\epsilon_s})$ and $e_{maxi} = max(e_{1i}, e_{2i}, e_{3i})$. In math terms this function is identical to the traditional one, however, it enables us to use large numbers and avoid the values computers fail to estimate, reducing the downward bias present in the optimization output.

Initial parameters: Yan and Zhang (2012) algorithm

Yang and Zhang (2012) finds that Likelihood maximization in PIN model often results into boundary solutions, which in turn creates bias. In order to avoid it, special algorithm was developed for choosing the initial parameters values, used in the optimization process. Firstly, they derive the marginal expected values of B (total buy trades)and S (total sell trades):

$$E(B) = \alpha(1-\delta)\mu + \epsilon_b$$

$$E(S) = \alpha\delta\mu + \epsilon_s$$

Using these two equations above, authors are able to set the initial values for parameters:

$$\alpha_0 = \alpha_i,\ \delta_0 = \delta_j,\ \epsilon_b^0 = \gamma_k \bar{B} \tag{2}$$

$$\mu_0 = \frac{\bar{B} - \epsilon_b^0}{\alpha_0(1-\delta_0)} \tag{3}$$

$$\epsilon_s^0 = \bar{S} - \alpha^0 \delta^0 \mu^0 \tag{4}$$

where α_i,δ_j,γ_k take one of equally-distanced values (0.1,0.3,0.5,0.7,0.9) at a time (there are 125 possible combinations), $\bar{B}$ and $\bar{S}$ are estimators of expected values of B and S respectively. Thus, we arrive at the following procedure: firstly, we run maximization for the 125 sets of the possible values, excluding those where ϵ_s is negative ($\epsilon_s < 0$). Secondly, if all solutions are on the boundary, we choose the one with the highest value of the Likelihood function, otherwise we exclude them and choose the one among non-boundary, using the same approach.

2 New Factorization technique and Initial parameters algorithm for Adjusted PIN

Using the intuition and ideas from Lin and Ke (2011) and Yang and Zhang (2012) models for PIN, in the following two subsections we derive their modified versions for Adjusted PIN framework.

New Factorization technique

As in Lin and Ke (2011), we base our derivation on two main ideas:

1) Computer provides more stable estimates for e^{x+y} rather than for $e^x e^y$

2) We should avoid plugging **too large** inputs into $exp()$ and **too low** ones into $ln()$. For instance, if we want to estimate $ln(e^{x+y} + e^z)$ we should better rewrite as:

$$ln\left(\frac{(e^{x+y} + e^z)e^k}{e^k}\right) = ln(e^{(x+y)-k} + e^{(z-k)}) + k$$

where $k = max(x + y,z)$

This trick guarantees the expression inside logarithm is always greater than one and we do not obtain e^x, $x > 710$, leading to overflow.

Applying these two principles on the initial Likelihood function, we get the more accurate expression (See **Appendix** for derivations):

$$\begin{aligned} L(\Theta|B_i,S_i) = & ln[(1-\alpha)(1-\theta)exp(-e_{maxi}) + (1-\alpha)\theta exp(e_{1i} - e_{maxi}) + \alpha(1-\theta')(1-\delta)exp(e_{2i} - e_{maxi}) + \\ & + \alpha\theta'(1-\delta exp(e_{3i} - e_{maxi}) + \alpha(1-\theta')\delta exp(e_{4i} - e_{maxi}) + \alpha\theta'\delta exp(e_{5i} - e_{maxi})] - \\ & - (\epsilon_b + \epsilon_s) + S_i ln(\epsilon_s) + B_i ln(\epsilon_b) + e_{maxi} - ln(B_i!S_i!) \end{aligned}$$

where $e_{1i} = -\Delta_b - \Delta_s + B_i ln(1 + \Delta_b/\epsilon_b) + S_i ln(1 + \Delta_s/\epsilon_s)$, $e_{2i} = -\mu_s + S_i ln(1 + \mu_s/\epsilon_s)$,$e_{3i} = -\mu_s - \Delta_b - \Delta_s + B_i ln(1 + \Delta_b/\epsilon_b) + S_i ln(1 + [\mu_s + \Delta_s]/\epsilon_s)$, $e_{4i} = -\mu_b + B_i ln(1 + \mu_b/\epsilon_b)$, $e_{5i} = -\mu_b - \Delta_b - \Delta_s + B_i ln(1 + [\mu_b + \Delta_b]/\epsilon_b) + S_i ln(1 + \Delta_s/\epsilon_s)$, $e_{maxi} = max(e_{1i}, e_{2i}, e_{3i}, e_{4i}, e_{5i})$.

New Initial Parameters Algorithm

We will use a more parsimonious specification with 10 parameters to estimate (setting $\theta = \theta'$):

$$\Theta = (\alpha, \delta, \theta, \mu_s, \mu_b, \epsilon_s, \epsilon_b, \Delta_s, \Delta_b)$$

We use 1-st and 2-nd moment conditions (See **Appendix** for derivations):

$$\begin{aligned}
\mathbb{E}(B) &= \epsilon_b + \theta\Delta_b + \alpha\delta\mu_b \\
\mathbb{E}(S) &= \epsilon_s + \theta\Delta_s + \alpha(1-\delta)\mu_s \\
\mathbb{E}(B^2) &= \epsilon_b^2 + \alpha\delta\mu_b^2 + \theta(\Delta_b^2 + 2\epsilon_b\Delta_b) + 2\alpha\delta\mu_b(\epsilon_b + \theta\Delta_b) \\
\mathbb{E}(S^2) &= \epsilon_s^2 + \alpha(1-\delta)\mu_s^2 + \theta(\Delta_s^2 + 2\epsilon_s\Delta_s) + 2\alpha(1-\delta)\mu_s(\epsilon_s + \theta\Delta_s)
\end{aligned}$$

As $\mathbb{E}(B)$ is always greater than ϵ_b, so we set the latter to be proportion of the sample analogue $\epsilon_b = \gamma\bar{B}$, where $\gamma = \{0.1, 0.3, 0.5, 0.7, 0.9\}$. In the same fashion we set $\epsilon_s = \gamma'\bar{S}$, where $\gamma' = \{0.1, 0.3, 0.5, 0.7, 0.9\}$.
As for probabilities of signal, positive signal and symmetric order flow shock we assign the following set of potential values to them:

$$\begin{aligned}
\alpha &= \{0.1, 0.3, 0.5, 0.7, 0.9\} \\
\delta &= \{0.1, 0.3, 0.5, 0.7, 0.9\} \\
\theta &= \{0.1, 0.3, 0.5, 0.7, 0.9\}
\end{aligned}$$

As a result, we have $5^5 = 3125$ initial values to estimate, which makes the algorithm extremely computationally intensive. Thus, we in further analysis and testing we set $\gamma = \gamma'$ which decreases the number of points to 625. Moreover, we will also lose some of them due to elimination of negative roots.

3 Data

We use common approach of testing the accuracy of PIN-type estimation algorithms, used in papers such as Gan et al (2015), etc. We generate 1000 datasets with 24 days in each (traditionally 60 is used, but as we deal with extremely computationally extensive framework, we use less observations to decrease estimation time). Via uniform distribution (U[0,1]) we generate $\alpha, \delta, \theta, p_\mu, p_{\mu_b}, p_\epsilon, p_{\epsilon_b}, p_{\Delta_b}$, which stand for probabilities of signal, positive signal, symmetric order flow, proportion of total trade intensity (I) which is generally informed, informed who buy, generally uninformed, uninformed who buy and symmetric who buy, respectively. We also set $I = 2500$. Thus, we get the following theoretical rates:

$$\begin{aligned}
\mu_b &= p_\mu \times p_{\mu_b} \times I \\
\mu_s &= p_\mu \times (1 - p_{\mu_b}) \times I \\
\epsilon_b &= (1 - p_\mu) \times p_\epsilon \times p_{\epsilon_b} \times I \\
\epsilon_s &= (1 - p_\mu) \times p_\epsilon \times (1 - p_{\epsilon_b}) \times I \\
\Delta_b &= (1 - p_\mu) \times (1 - p_\epsilon) \times p_{\Delta_b} \times I \\
\Delta_s &= (1 - p_\mu) \times (1 - p_\epsilon) \times (1 - p_{\Delta_b}) \times I
\end{aligned}$$

Finally, we generate, using Poisson distribution, the Buy and Sell trading rates, which we further use to obtain the empirical estimations of Adjusted PIN.

4 Testing Efficiency of Algorithms

We consider four specifications, which combine in different ways new initial parameters and likelihood specifications, proposed by this paper, and their standard versions:

(1) Duarte Likelihood + Simple Initial Parameters algorithm
(2) Duarte Likelihood + New Initial Parameters algorithm
(3) New Likelihood + Simple Initial Parameters algorithm
(4) New Likelihood + New Initial Parameters algorithm

By "Simple Initial Parameters algorithm" we mean taking 10 different starting points, the approach used by Duarte et al (2007). We apply each of these specifications above on the order flow, generated using the methodology described in the previous section, and compare to the theoretical values of parameters and of Adjusted PIN.

4.1 Estimation Speed and Accuracy

Taking order flow vector of 24 estimates, we arrive at the following timings:

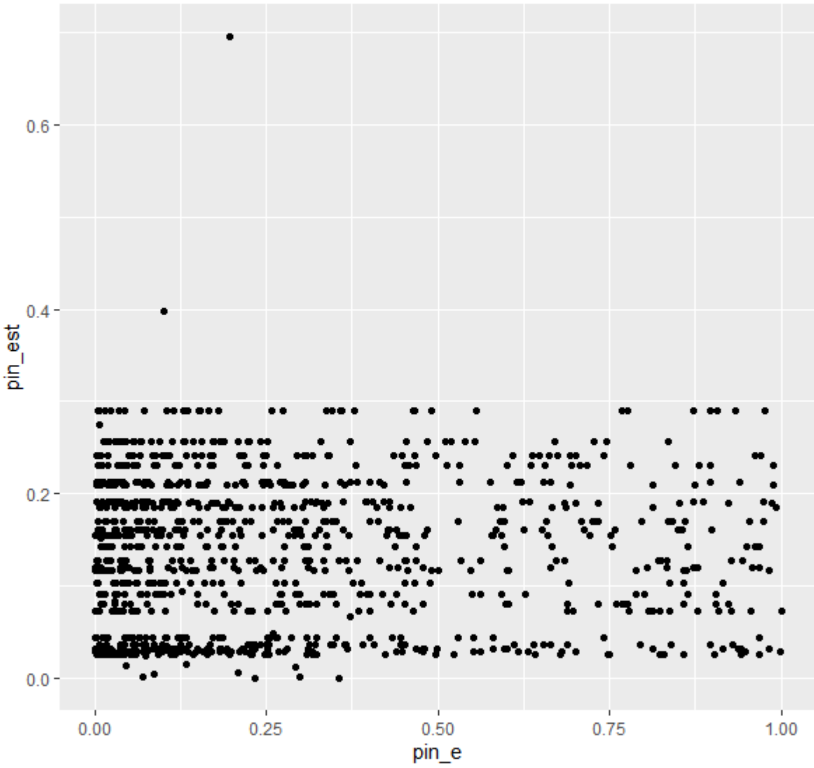
Table 1: Time to estimate

Model Specification	Time
(1) Duarte Likelihood + Simple Initial Parameters algorithm	3.317 secs
(2) Duarte Likelihood + New Initial Parameters algorithm	113.331 secs
(3) New Likelihood + Simple Initial Parameters algorithm	0.82728 secs
(4) New Likelihood + New Initial Parameters algorithm	15.46419 secs

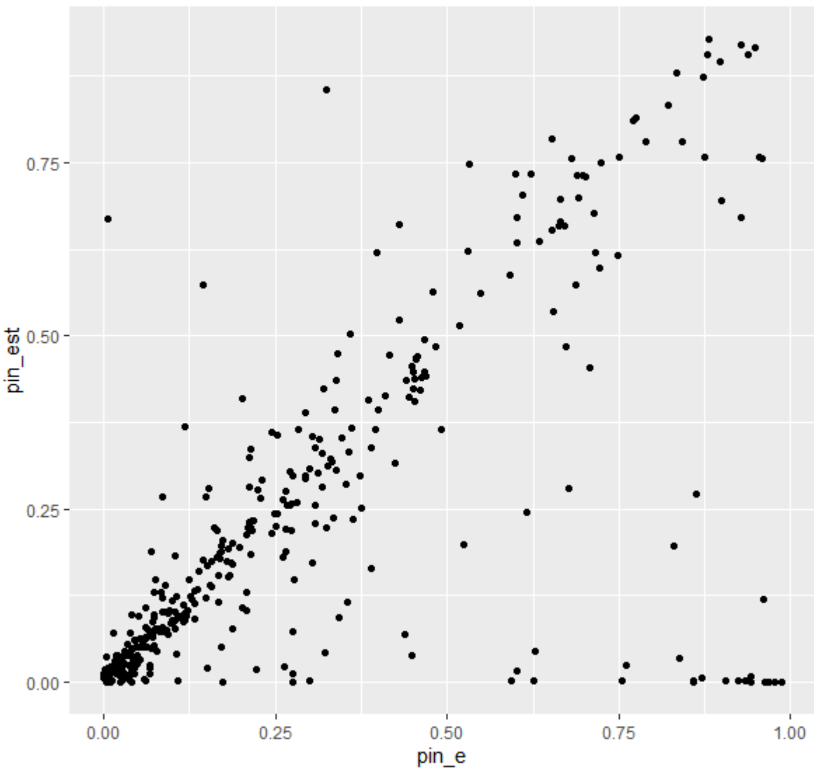
The modified version of likelihood significantly increases the speed of estimation several times, while the new initial parameters algorithm, as expected, extremely slows down the calculation process, thus, the question of accuracy and time estimation trade-off is vital.

Table 2: Mean Squared Error

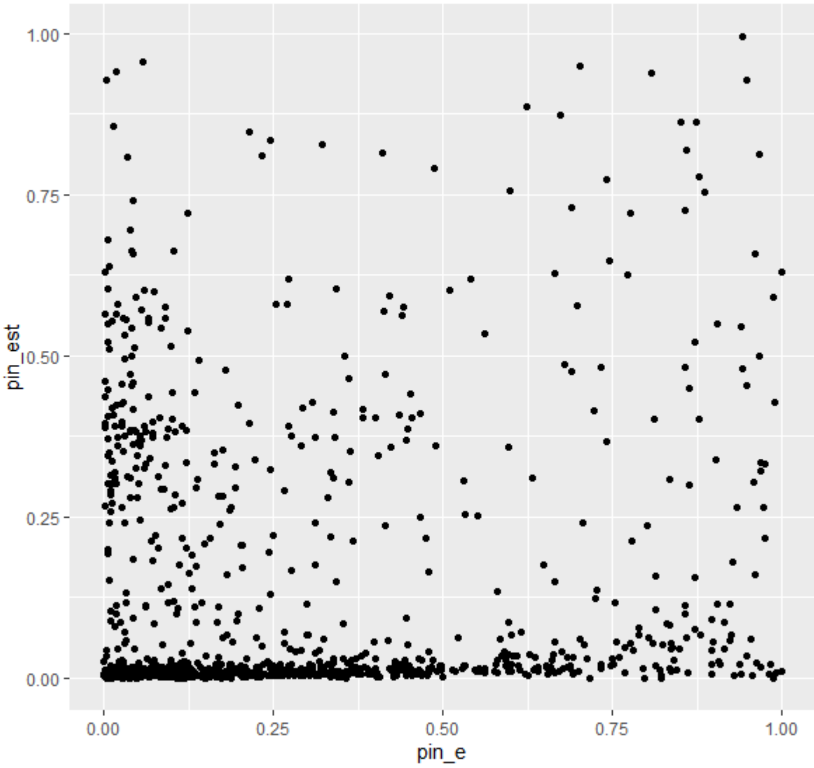
Model Specification	MSE
(1) Duarte Likelihood + Simple Initial Parameters algorithm	0.1066
(2) Duarte Likelihood + New Initial Parameters algorithm	0.0493
(3) New Likelihood + Simple Initial Parameters algorithm	0.0704
(4) New Likelihood + New Initial Parameters algorithm	0.0463



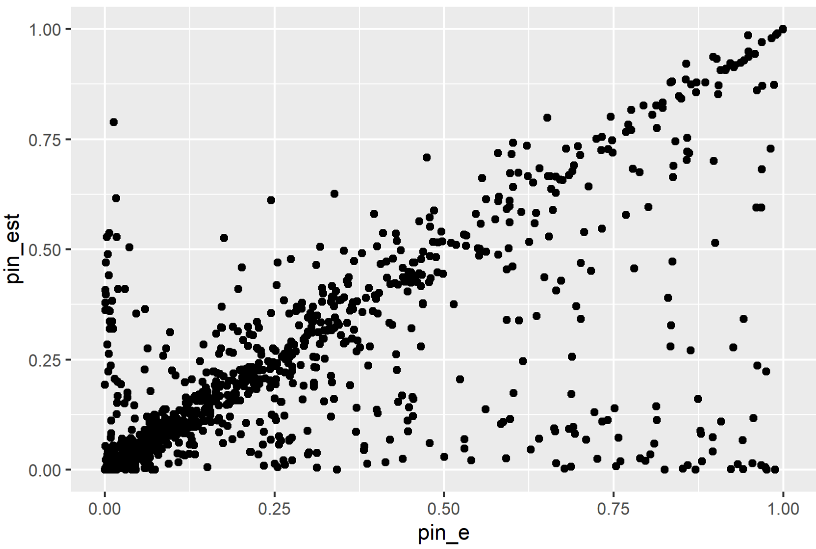
Duarte Likelihood + Simple Initial Parameters algorithm **(PIN)**



Duarte Likelihood + New Initial Parameters algorithm **(PIN)**

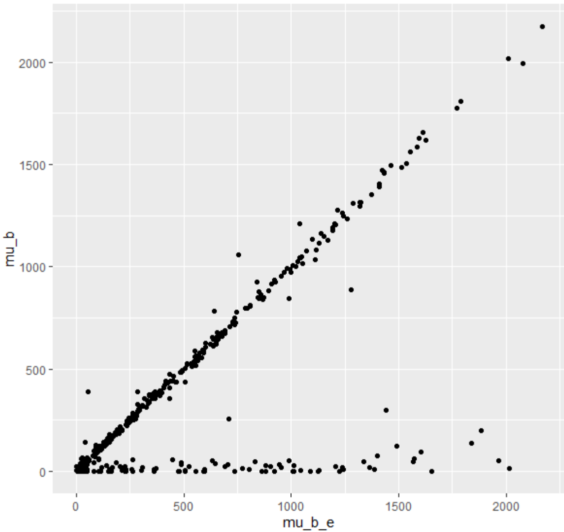


New Likelihood + Simple Initial Parameters algorithm **(PIN)**

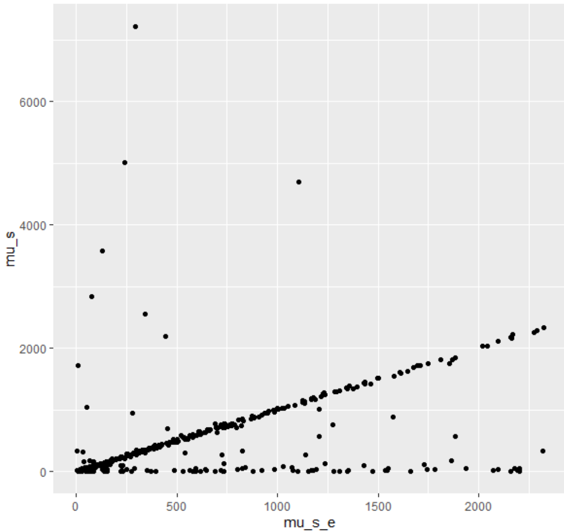


New Likelihood + New Initial Parameters algorithm **(PIN)**

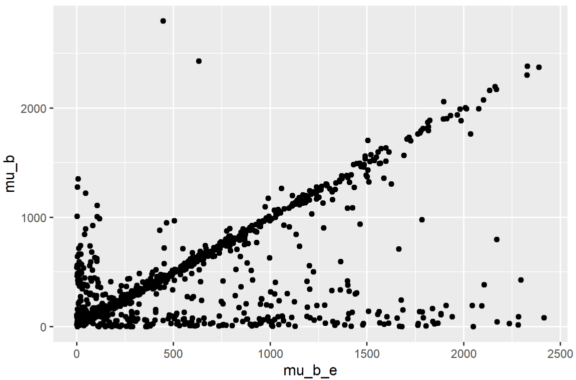
Comparing specifications (1) vs (2) and (3) vs (4), it is clear that New likelihood and New initial parameters algorithm each separately improve the accuracy of estimations. Modified likelihood also significantly increases the speed of estimations. However, the new initial parameters algorithm seems to improve the efficiency to higher extent in case of Duarte Likelihood [(1) -> (2)], compared to the modified one [(3) -> (4)]. Surprisingly, (2) provides almost the same in terms of accuracy results as (4) and both specifications outperform all the rest models in predicting μ_b and μ_s, However, (2) is ten times slower to be estimated, which makes specification (4) the most efficient one in terms of time vs efficiency trade-off.



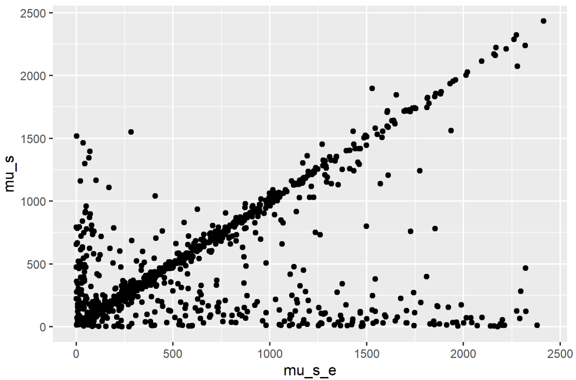
Model (2) (μ_b)



Model (2) (μ_b)



Model (4) (μ_b)



Model (4) (μ_s)

5 Conclusion

This paper introduces two modifications for estimation procedure of Adjusted PIN model. Being the nested model of the traditional PIN by Easley et al (1996), it inherits the **FPE** problem, which we effectively solve by providing the new version of Likelihood and increase the computation speed, based on the ideas and intuition of Lin and Ke (2011). In order to enhance the efficiency of optimization and avoid **false boundary solutions**, we introduce the new initial parameters algorithm, using method of moments and iteration. Thus, these two new approaches above make the Adjusted PIN estimates much more efficient and, thus, comparable with traditional PIN outcomes with Lin and Ke (2011) and Yang and Zhang (2012) modifications.

6 Bibiliography

[1] Duarte, Jefferson, and Young L. (2009), Why is PIN priced?, Journal of Financial Economics 91, 119—138.

[2] Easley D., Hvidkjaer S., and O'Hara M. (2010), Factoring information into returns. Journal of Financial and Quantitative Analysis.

[3] Easley D., Kiefer N., O'Hara M., and Paperman J. (1996), Liquidity, information, and infrequently traded stocks, Journal of Finance 51, 1405- 1436.

[4] Easley D., O'Hara M. (1987), Price, trade size, and information in securities markets, Journal of Financial Economics 19, 69-90.

[5] Gan Q., Wei W.C., Johnstone D. (2015) A faster estimation method for the probability of informed trading using hierarchical agglomerative clustering. Quantitative Finance, 15(11):1805–1821.

[6] Glosten L. and P. Milgrom (1985), Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, Journal of Financial Economics 13, 71-100.

[7] G. Kuzmin (2021)

[8] H.-W. W. Lin and W.-C. Ke (2011), A computing bias in estimating the probability of informed trading. Journal of Financial Markets, 14(4), 625–640.

[9] Yan, Y. and S. Zhang (2012), An improved estimation method and empirical properties of the probability of informed trading, Journal of Banking Finance 36, 454-467

Appendix

Factorization Derivation

The initial likelihood from Duarte et al (2007) has the following form:

$$
\begin{aligned}
L(\Theta|B_i,S_i) =ln \Bigg[&(1-\alpha)(1-\theta)exp(-\epsilon_b-\epsilon_s)\frac{\epsilon_b^{B_i}\epsilon_s^{S_i}}{B_i!S_i!}+ \\
&+(1-\alpha)\theta exp(-\epsilon_b-\epsilon_s-\Delta_b-\Delta_s)\frac{(\epsilon_b+\Delta_b)^{B_i}(\epsilon_s+\Delta_s)^{S_i}}{B_i!S_i!}+ \\
&+\alpha(1-\theta')(1-\delta)exp(-\epsilon_b-\mu_s-\epsilon_s)\frac{\epsilon_b^{B_i}(\mu_s+\epsilon_s)^{S_i}}{B_i!S_i!}+ \\
&+\alpha\theta'(1-\delta)exp(-\epsilon_b-\epsilon_s-\mu_s-\Delta_b-\Delta_s)\frac{(\epsilon_b+\Delta_b)^{B_i}(\mu_s+\epsilon_s+\Delta_s)^{S_i}}{B_i!S_i!}+ \\
&+\alpha(1-\theta')\delta exp(-\mu_b-\epsilon_b-\epsilon_s)\frac{(\mu_b+\epsilon_b)^{B_i}\epsilon_s^{S_i}}{B_i!S_i!}+ \\
&+\alpha\theta'\delta exp(-\mu_b-\epsilon_b-\epsilon_s-\Delta_b-\Delta_s)\frac{(\mu_b+\epsilon_b+\Delta_b)^{B_i}(\epsilon_s+\Delta_s)^{S_i}}{B_i!S_i!}\Bigg]
\end{aligned}
$$

Taking out common terms:

$$
\begin{aligned}
L(\Theta|B_i,S_i) =ln\{ &\frac{1}{B_i!S_i!}exp(-\epsilon_b-\epsilon_s)[(1-\alpha)(1-\theta)\epsilon_b^{B_i}\epsilon_s^{S_i}+ \\
&+(1-\alpha)\theta exp(-\Delta_b-\Delta_s)(\epsilon_b+\Delta_b)^{B_i}(\epsilon_s+\Delta_s)^{S_i}+ \\
&+\alpha(1-\theta')(1-\delta)exp(-\mu_s)\epsilon_b^{B_i}(\mu_s+\epsilon_s)^{S_i}+ \\
&+\alpha\theta'(1-\delta)exp(-\mu_s-\Delta_b-\Delta_s)(\epsilon_b+\Delta_b)^{B_i}(\mu_s+\epsilon_s+\Delta_s)^{S_i}+ \\
&+\alpha(1-\theta')\delta exp(-\mu_b)(\mu_b+\epsilon_b)^{B_i}\epsilon_s^{S_i}+ \\
&+\alpha\theta'\delta exp(-\mu_b-\Delta_b-\Delta_s)(\mu_b+\epsilon_b+\Delta_b)^{B_i}(\epsilon_s+\Delta_s)^{S_i}\}
\end{aligned}
$$

Divide and multiply inside $ln()$ by $\epsilon_b^{B_i}\epsilon_s^{S_i}$:

$$\begin{aligned}
L(\Theta|B_i,S_i) =& ln[(1-\alpha)(1-\theta)+ \\
&+(1-\alpha)\theta exp(-\Delta_b-\Delta_s)\left(\frac{\epsilon_b+\Delta_b}{\epsilon_b}\right)^{B_i}\left(\frac{\epsilon_s+\Delta_s}{\epsilon_s}\right)^{S_i}+ \\
&+\alpha(1-\theta')(1-\delta)exp(-\mu_s)\left(\frac{\mu_s+\epsilon_s}{\epsilon_s}\right)^{S_i}+ \\
&+\alpha\theta'(1-\delta)exp(-\mu_s-\Delta_b-\Delta_s)\left(\frac{\epsilon_b+\Delta_b}{\epsilon_b}\right)^{B_i}\left(\frac{\mu_s+\epsilon_s+\Delta_s}{\epsilon_s}\right)^{S_i}+ \\
&+\alpha(1-\theta')\delta exp(-\mu_b)\left(\frac{\mu_b+\epsilon_b}{\epsilon_b}\right)^{B_i}+ \\
&+\alpha\theta'\delta exp(-\mu_b-\Delta_b-\Delta_s)\left(\frac{\mu_b+\epsilon_b+\Delta_b}{\epsilon_b}\right)^{B_i}\left(\frac{\epsilon_s+\Delta_s}{\epsilon_s}\right)^{S_i}]- \\
&-(\epsilon_b+\epsilon_s)+S_i ln(\epsilon_s)+B_i ln(\epsilon_b)-ln(B_i!S_i!)
\end{aligned}$$

Several more simplifications and transformations:

$$\begin{aligned}
L(\Theta|B_i,S_i) =& ln[(1-\alpha)(1-\theta)+ \\
&+(1-\alpha)\theta exp\left(-\Delta_b-\Delta_s+B_i ln\left(1+\frac{\Delta_b}{\epsilon_b}\right)+S_i ln\left(1+\frac{\Delta_s}{\epsilon_s}\right)\right)+ \\
&+\alpha(1-\theta')(1-\delta)exp\left(-\mu_s+S_i ln\left(1+\frac{\mu_s}{\epsilon_s}\right)\right)+ \\
&+\alpha\theta'(1-\delta)exp\left(-\mu_s-\Delta_b-\Delta_s+B_i ln\left(1+\frac{\Delta_b}{\epsilon_b}\right)+S_i ln\left(1+\frac{\mu_s+\Delta_s}{\epsilon_s}\right)\right)+ \\
&+\alpha(1-\theta')\delta exp\left(-\mu_b+B_i ln\left(1+\frac{\mu_b}{\epsilon_b}\right)\right)+ \\
&+\alpha\theta'\delta exp\left(-\mu_b-\Delta_b-\Delta_s+B_i ln\left(1+\frac{\mu_b+\Delta_b}{\epsilon_b}\right)+S_i ln\left(1+\frac{\Delta_s}{\epsilon_s}\right)\right)]- \\
&-(\epsilon_b+\epsilon_s)+S_i ln(\epsilon_s)+B_i ln(\epsilon_b)-ln(B_i!S_i!)
\end{aligned}$$

We can drop $ln(B_i!S_i!)$ and define the subparts of the Likelihood above as:

$$e_{1i} = -\Delta_b - \Delta_s + B_i ln(1 + \Delta_b/\epsilon_b) + S_i ln(1 + \Delta_s/\epsilon_s)$$

$$e_{2i} = -\mu_s + S_i ln(1 + \mu_s/\epsilon_s)$$

$$e_{3i} = -\mu_s - \Delta_b - \Delta_s + B_i ln(1 + \Delta_b/\epsilon_b) + S_i ln(1 + [\mu_s + \Delta_s]/\epsilon_s)$$

$$e_{4i} = -\mu_b + B_i ln(1 + \mu_b/\epsilon_b)$$

$$e_{5i} = -\mu_b - \Delta_b - \Delta_s + B_i ln(1 + [\mu_b + \Delta_b]/\epsilon_b) + S_i ln(1 + \Delta_s/\epsilon_s)$$

Thus, we have:

$$L(\Theta|B_i,S_i) = ln[(1-\alpha)(1-\theta) + (1-\alpha)\theta exp(e_{1i}) + \alpha(1-\theta')(1-\delta)exp(e_{2i}) + \alpha\theta'(1-\delta)exp(e_{3i}) +$$
$$+ \alpha(1-\theta')\delta exp(e_{4i}) + \alpha\theta'\delta exp(e_{5i})] - (\epsilon_b + \epsilon_s) + S_i ln(\epsilon_s) + B_i ln(\epsilon_b) - ln(B_i!S_i!)$$

As we have already mentioned in "New Factorization technique" section, to estimate expressions such as $ln(e^{x+y} + e^z)$, we should better rewrite as:

$$ln\left(\frac{(e^{x+y} + e^z)e^k}{e^k}\right) = ln(e^{(x+y)-k} + e^{(z-k)}) + k$$

where $k = max(x+y,z)$

This trick helps to avoid **FPE** problem

Defining $e_{maxi} = max(e_{1i}, e_{2i}, e_{3i}, e_{4i}, e_{5i})$, we perform the final transformation as in the example above and obtain the New Likelihood expression:

$$L(\Theta|B_i,S_i) = ln[(1-\alpha)(1-\theta)exp(-e_{maxi}) + (1-\alpha)\theta exp(e_{1i} - e_{maxi}) + \alpha(1-\theta')(1-\delta)exp(e_{2i} - e_{maxi}) +$$
$$+ \alpha\theta'(1 - \delta exp(e_{3i} - e_{maxi}) + \alpha(1-\theta')\delta exp(e_{4i} - e_{maxi}) + \alpha\theta'\delta exp(e_{5i} - e_{maxi})] -$$
$$-(\epsilon_b + \epsilon_s) + S_i ln(\epsilon_s) + B_i ln(\epsilon_b) + e_{maxi} - ln(B_i!S_i!)$$

Initial Parameters moments derivation

Adjusted PIN Likelihood:

$$
\begin{aligned}
P(B{,}S) = & ln \Bigg[(1-\alpha)(1-\theta)exp(-\epsilon_b)\frac{\epsilon_b^B}{B!}exp(-\epsilon_s)\frac{\epsilon_s^S}{S!}+ \\
& + (1-\alpha)\theta exp(-\epsilon_b-\Delta_b)\frac{(\epsilon_b+\Delta_b)^B}{B!}exp(-\epsilon_s-\Delta_s)\frac{(\epsilon_s+\Delta_s)^S}{S!}+ \\
& + \alpha(1-\theta')(1-\delta)exp(-\epsilon_b)\frac{\epsilon_b^B}{B!}exp(-\mu_s-\epsilon_s)\frac{(\mu_s+\epsilon_s)^S}{S!}+ \\
& + \alpha\theta'(1-\delta)exp(-\epsilon_b-\Delta_b)\frac{(\epsilon_b+\Delta_b)^B}{B!}exp(-\epsilon_s-\mu_s-\Delta_s)\frac{(\mu_s+\epsilon_s+\Delta_s)^S}{S!}+ \\
& + \alpha(1-\theta')\delta exp(-\mu_b-\epsilon_b)\frac{(\mu_b+\epsilon_b)^B}{B!}exp(-\epsilon_s\frac{(\epsilon_s^S}{S!})+ \\
& + \alpha\theta'\delta exp(-\mu_b-\epsilon_b-\Delta_b)\frac{(\mu_b+\epsilon_b+\Delta_b)^B}{B!}exp(-\epsilon_s-\Delta_s)\frac{(\epsilon_s+\Delta_s)^S}{S!} \Bigg]
\end{aligned}
$$

The marginal probability for the Buy order:

$$P(B)=\sum_{S=0}^{\infty}P(B,S)=ln\Bigg[(1-\alpha)(1-\theta)exp(-\epsilon_b)\frac{\epsilon_b^B}{B!}exp(-\epsilon_s)\sum_{S=0}^{\infty}\frac{\epsilon_s^S}{S!}+$$

$$+(1-\alpha)\theta exp(-\epsilon_b-\Delta_b)\frac{(\epsilon_b+\Delta_b)^B}{B!}exp(-\epsilon_s-\Delta_s)\sum_{S=0}^{\infty}\frac{(\epsilon_s+\Delta_s)^S}{S!}+$$

$$+\alpha(1-\theta')(1-\delta)exp(-\epsilon_b)\frac{\epsilon_b^B}{B!}exp(-\mu_s-\epsilon_s)\sum_{S=0}^{\infty}\frac{(\mu_s+\epsilon_s)^S}{S!}+$$

$$+\alpha\theta'(1-\delta)exp(-\epsilon_b-\Delta_b)\frac{(\epsilon_b+\Delta_b)^B}{B!}exp(-\epsilon_s-\mu_s-\Delta_s)\sum_{S=0}^{\infty}\frac{(\mu_s+\epsilon_s+\Delta_s)^S}{S!}+$$

$$+\alpha(1-\theta')\delta exp(-\mu_b-\epsilon_b)\frac{(\mu_b+\epsilon_b)^B}{B!}exp(-\epsilon_s)\sum_{S=0}^{\infty}(\frac{\epsilon_s^S}{S!})+$$

$$+\alpha\theta'\delta exp(-\mu_b-\epsilon_b-\Delta_b)\frac{(\mu_b+\epsilon_b+\Delta_b)^B}{B!}exp(-\epsilon_s-\Delta_s)\sum_{S=0}^{\infty}\frac{(\epsilon_s+\Delta_s)^S}{S!}\Bigg]$$

Using Taylor series $(e^x=\sum^{\infty}\frac{x^n}{n!})$**, we obtain:**

$$P(B)=ln\Bigg[(1-\alpha)(1-\theta)exp(-\epsilon_b)\frac{\epsilon_b^B}{B!}+(1-\alpha)\theta exp(-\epsilon_b-\Delta_b)\frac{(\epsilon_b+\Delta_b)^B}{B!}+$$

$$+\alpha(1-\theta')(1-\delta)exp(-\epsilon_b)\frac{\epsilon_b^B}{B!}+\alpha\theta'(1-\delta)exp(-\epsilon_b-\Delta_b)\frac{(\epsilon_b+\Delta_b)^B}{B!}+$$

$$\alpha(1-\theta')\delta exp(-\mu_b-\epsilon_b)\frac{(\mu_b+\epsilon_b)^B}{B!}+\alpha\theta'\delta exp(-\mu_b-\epsilon_b-\Delta_b)\frac{(\mu_b+\epsilon_b+\Delta_b)^B}{B!}\Bigg]$$

Using the same Taylor transformation, we estimate the Expected number of

Buy orders as:

$$\mathbb{E}(B) = \sum_{B=0}^{\infty} P(B) \times B = (1-\alpha)(1-\theta)exp(-\epsilon_b)\frac{\epsilon_b^{B-1}\epsilon_b}{(B-1)!} +$$

$$+ (1-\alpha)\theta exp(-\epsilon_b - \Delta_b)\frac{(\epsilon_b + \Delta_b)^{B-1}(\epsilon_b + \Delta_b)}{(B-1)!} + ... =$$

$$= \epsilon_b + (1-\alpha)\theta\Delta_b + \alpha\theta'\Delta_b + \alpha\delta\mu_b$$

Using the same approach, we can derive the Expected number of Sell orders:

$$\mathbb{E}(S) = \sum_{S=0}^{\infty} P(S) \times S = \epsilon_s + (1-\alpha)\theta\Delta_s + \alpha\theta'\Delta_S + \alpha(1-\delta)\mu_s$$

Deriving the second moments:

$$\mathbb{E}(B^2) = \sum_{B=0}^{\infty} P(B) \times B^2 = (1-\alpha)(1-\theta)exp(-\epsilon_b)\frac{\epsilon_b^{B-2}\epsilon_b}{(B-2)!} + ... =$$

$$= \epsilon_b^2 + \alpha\delta\mu_b^2 + ((1-\alpha)\theta + \alpha\theta')(\Delta_b^2 + 2\epsilon_b\Delta_b) + 2\alpha\delta\mu_b(\epsilon_b + \theta'\Delta_b)$$

$$\mathbb{E}(S^2) = \sum_{S=0}^{\infty} P(S) \times S^2 = \epsilon_s^2 + \alpha(1-\delta)\mu_s^2 + ((1-\alpha)\theta + \alpha\theta')(\Delta_s^2 + 2\epsilon_s\Delta_s) + 2\alpha\theta'(1-\delta)\Delta_s\mu_s +$$

$$+ 2\alpha(1-\delta)\epsilon_s\mu_s$$

Thus, setting $\theta' = \theta$, we have the following system of moment conditions:

$$\mathbb{E}(B) = \epsilon_b + \theta\Delta_b + \alpha\delta\mu_b$$

$$\mathbb{E}(S) = \epsilon_s + \theta\Delta_s + \alpha(1-\delta)\mu_s$$

$$\mathbb{E}(B^2) = \epsilon_b^2 + \alpha\delta\mu_b^2 + \theta(\Delta_b^2 + 2\epsilon_b\Delta_b) + 2\alpha\delta\mu_b(\epsilon_b + \theta\Delta_b)$$

$$\mathbb{E}(S^2) = \epsilon_s^2 + \alpha(1-\delta)\mu_s^2 + \theta(\Delta_s^2 + 2\epsilon_s\Delta_s) + 2\alpha(1-\delta)\mu_s(\epsilon_s + \theta\Delta_s)$$